Claims

We claim:

1   1. A system for encoding a plurality of videos acquired of a moving object in a

2   scene by a plurality of fixed cameras, comprising:

3   means for determining camera calibration data of each camera;

4   means for associating the camera calibration data of each camera with the

5   video acquired by the camera;

6   means for determining a segmentation mask for each frame of each video,

7   the segmentation mask identifying only pixels in the frame associated with the

8   object;

9   a shape encoder configured to encode the segmentation masks;

10   a position encoder configured to encode a position of each pixel; and

11   a color encoder configured to encode a color of each pixel.

1   2. The system of claim 1, further comprising:

2   a multiplexer configured to combine outputs of the shape encoder, the

3   position encoder, and the color encoder into a single bitstream.

1   3. The system of claim 2, further comprising:

2   a decoder;

3   means for transferring the bitstream to the decoder; and

4   rendering a decoded bitstream from an arbitrary viewpoint using the camera

5   calibration data.

1   4. The system of claim 3, in which the arbitrary viewpoint is constrained in space.

1   5. The system of claim 3, in which the arbitrary viewpoint is unconstrained in

2   space.

1   6. The system of claim 1, further comprising:

2   means for maintaining a dynamic 3D point model defining a geometry of the

3   moving object.

1   7. The system of claim 1, in which each point of the dynamic 3D point model is

2   associated with an identifier of one or more of the plurality of cameras.

1   8. The system of claim 1, in which the encoded segmentation masks are

2   compressed using a lossless compression, and the position and the colors are

3   encoded using a lossy compression.

1   9. The method of claim 1, in which the camera calibration data are updated

2   periodically when any of the fixed cameras are relocated.

1   10. The system of claim 1, in which the segmentation masks are encoded using

2   MPEG-4 lossless binary shape encoding, the positions include depth values

3   encoded as quantized pixel luminance values, and the colors are encoded using

4   MPEG-4 video object coding.

1   11. The method of claim 1, in which the entire scene is encoded using a scene

2   specifying relations between static and dynamic portions of the scene.

1   12. The system of claim 1, further comprising:

2   a decoder configured to decode the encoded segmentation masks, the

3   encoded positions, and the encoded colors as an output video having an arbitrary

4   viewpoint using the camera calibration data.

1   13. The system of claim 12, in which the arbitrary viewpoint is different than a

2   viewpoint of any of the cameras.

1   14. The system of claim 12, in which images of the output video are composited

2   with a virtual scene.

1   15. The system of claim 12, in which a playback frame rate of the output video is

2   different than a frame rate used to acquired the videos by the plurality of cameras.

1.   16. The system of claim 8, in which the lossy compression scheme is a progressive

2   encoding using embedded zerotree wavelet coding.

1   17. The system of claim 1, in which the shape encoder use MPEG-4 lossless binary

2   shape encoding, the position encoder encodes depth values, and the color encoder

3   uses MPEG-4 video object coding.

1   18. The system of claim 1, further comprising:

2   means for partitioning each video into a plurality of segments, each segment

3   including a plurality of frames; and

4   means for encoding a key frame and difference frames of each segment,

5   using the shape encoder, the position encoder, and the color encoder into a single

6   bitstream.

32

1   19. The system of claim 18, in which the key frames comprise a base layer of an

2   encoded video bitstream, and the difference frames comprise an enhancement layer

3   of the encoded bitstream.

1   20. The system of claim 18, further comprising:

2   means for averaging the frames in each segment to construct the key frame;

3   means for determining the difference frame for each frame in the segment

4   from the key frame and the frame.

1   21. The system of claim 18, in which the key frame is a first frame of the segment,

2   and a difference frame is determined from a current frame and previous frames in.

3   the segment.

1   22. The system of claim 1, further comprising:

2   a surface normal encoder configured to encode a surface normal of each

3   pixel; and

4   a splat size encoder configured to encode a splat size for each pixel; and

5   means for combining the outputs of the surface normal encoder and the splat

6   size encoder with the single bitstream.

1   23. The system of claim 22, in which the surface normal vectors are progressively

2   encoded using an octahedron subdivision of a unit sphere and the splat sizes are

3   encoded as quantized codewords represented in a gray scale MPEG video object.

1   24. The system of claim 12, in which splat sizes and surface normals are estimated

2   from the positions.